Source: *The Straits Times,* pA21
Date: 12 August 2021

While it is hard to formally define what machine learning is exactly due to its interdisciplinary nature, one could say that it is a field that develops algorithms designed to be applied to data sets, with the main focus being prediction or clustering tasks.
PHOTO: ISTOCKPHOTO

Ask NUS economists

Machine learning – a powerful tool in economics

Machine learning algorithms can beat standard methods at macroeconomic forecasting. But the 'garbage in, garbage out' rule still applies.

Denis Tkachenko

For The Straits Times

Q *Do economists use machine learning methods and, if so, how are they useful for solving real-world economic problems?*

A Our lives are full of daily interactions with services powered by machine learning algorithms, be it avoiding traffic using Google Maps, looking through personalised recommendations on Netflix, humming a tune to music-identifying app Shazam or navigating your spam folder.

The term "machine learning" is perhaps among the biggest buzzwords that permeate the media in recent years. Oftentimes sensationally worded headlines make it sound like machines can learn just like humans do. Big success stories, predominantly in tech, are emphasised, such as DeepMind's AlphaGo defeating Go grandmaster Lee Sedol, and the progress seen in self-driving cars.

The question then is: What are these methods and can they be leveraged to guide economic decisions that improve the functioning of our society?

When hearing about machine learning in economics, one's vivid mind may draw a picture of a computer scanning economics textbooks and somehow processing them to "learn".

But what does machine learning entail in practice? The term itself goes back to the 1959 article by computer scientist Arthur Samuel in the IBM Journal of Research and Development, and its popularity has much to do with IBM marketing the term to win over clients and attract talented employees.

While it is hard to formally define what machine learning is exactly due to its interdisciplinary nature, one could say that it is a field that develops algorithms designed to be applied to data sets, with the main focus being prediction or clustering tasks.

To put it simply, it is like a chair with three legs: statistics, computer science and domain knowledge – in our case, economics.

As follows from above, machine learning pursues two broad types of tasks: prediction and grouping/summarising data sets. These tasks define two large subfields of machine learning – supervised and unsupervised learning, respectively. Despite the latter being useful for data visualisation and summary, it is the supervised learning methods that are of most interest to economists.

Supervised learning's name derives from the fact that one can imagine the algorithm as being a student who does his homework based on the teaching material: a data set consisting of inputs and outcomes.

The student's task is to come up with a way to combine data on available input variables to generate a rule to predict the outcome. For example, one may think of predicting whether a loan application will be approved based on the characteristics of the borrower, such as age, income, number of credit cards, and other loans outstanding.

The student then shows his answer to the teacher, who provides the correct answer and gives a "grade": some measure of predictive accuracy. Upon receiving the grade, the student goes back to modify his input-output model and the process goes on until the teacher is satisfied with the answer quality.

This is akin to learning by example, and the hope is that by the end of this process called training, we obtain a prediction rule that will be useful for the new inputs we are likely to encounter in practice. Note, however, that the student here is a rote learner and is unable to "think" creatively.

Machine learning methods excel at prediction and economists are asked to make predictions on a regular basis: Policymakers and businesses alike require reliable forecasts of economic activity, inflation, exchange rates, consumer confidence, to name but a few.

While established approaches tend to work reasonably well in general, the Covid-19 pandemic has made macroeconomic forecasting very challenging: The issue is not so much that the start of the pandemic was impossible to forecast using standard models, but that the economic data that comes afterwards is so extreme that it "throws off" further forecasts.

A recent study pitting traditional methods against machine learning alternatives using data from the United Kingdom has produced a surprising result. The machine learning algorithms overwhelmingly outperformed in short-term forecasting during the first year of the pandemic and were able to forecast the substantial "bounce back" after the initial shock that actually occurred, while the majority of models typically used in central banks wrongly projected a catastrophic worsening of economic conditions throughout 2020.

The machine learning forecasts still performed at least as well as standard models in "quieter times", thus demonstrating their all-around dominance.

Regular readers of this column may ask: "But what about program evaluation? Don't economists often deal with estimating causal effects of policies?"

Progress can be made here with machine learning as well. First, building a non-causal predictive model can still be useful since one can examine the variables that are most important for making correct predictions: For example, economists at the Ministry of Trade and Industry used this approach to discover that skills mismatch is key in explaining labour market outcomes, and an immediate policy implication is that more information on jobs posted online may reduce search frictions for job seekers.

Second, it was found that many causal effect estimation problems in economics can be represented as two separate machine learning problems, and the resulting approach, dubbed "double machine learning", allows for more accurate policy analysis.

To summarise, machine learning methods add a powerful tool to the arsenal of applied economists. Going forward, the integration of machine learning into economic practice will only grow and become one of the critical skills in the digital economy based on data-driven decision-making.

Universities in Singapore are recognising this by piloting interdisciplinary economics and data science programmes.

However, it is important to remember that machine learning alone is no silver bullet: The machine can learn only from the inputs provided by the researcher, so "garbage in, garbage out" still applies.

The researcher must carefully select the method and the data that are appropriate in the situation via domain expertise.

stopinion@sph.com.sg

• Denis Tkachenko is a senior lecturer at the Department of Economics, National University of Singapore.